

NIEHS Data Sharing Efforts-RFI and Workshop

**Kim McAllister, Kim Gray, Symma Finn, Cindy
Lawler, Caroline Dilworth, Jennifer Collins,
Claudia Thompson, Gwen Collman**

**National Institute of Environmental Health
Sciences**



Purpose:

- ❑ Explore unique considerations for data sharing for studies with environmental exposure data
- ❑ Examine challenges that exist for environmental health researchers wishing to more broadly share their data with others
- ❑ Identify best practices and successful models of data sharing that are applicable to environmental health data sharing
- ❑ Consider additional tools/resources that researchers believe will allow more efficient and effective data sharing in the environmental health science community

Unique Considerations for Sharing of Environmental Data

- The heterogeneity of environmental and biological measurements
- The potential to identify individuals based on the association of environmental exposures with geographical data
- The increased interest in return of individual or community-level research results from environmental research
- The regulatory implications of the use of environmental exposure and health data in developing US national research policies
- The unique concerns of vulnerable populations who are disproportionately impacted by environmental exposures.

Broad Themes Emphasized in RFI and Workshop Responses/Discussions:

- ☐ **Protection of Privacy/Confidentiality Issues**
- ☐ **Institutional Review Board (IRB) issues**
- ☐ **Legal and Regulatory Issues**
- ☐ **NIH Programmatic and Logistical Considerations**
- ☐ **Computational Challenges**

Protection of Privacy and Confidentiality:

- ❖ Importance of community-based participatory research, especially in vulnerable communities disproportionately affected by exposures
- ❖ Concerns regarding specific location identification of exposures and potential harms (insurance, employment, property values, reporting of regulated substances, etc.)
- ❖ “Anonomize” data:
- ❖ Security is inadequate for online databases

Institutional Review Boards (IRB):

The lack of continuity, consistency, and clarity across IRBs was emphasized by many research groups as a disincentive for attempting to more broadly share their data with others.

- ❑ Informed consent models need to be redeveloped to allow more sharing
- ❑ Education of unique considerations related to data sharing for environmental health data at IRBs is key.

Legal and Regulatory Concerns:

Exposure data will continue to be of high interest to regulatory agencies with respect to the evaluation of the health implications of chemicals:

- ☐ The reanalysis and/or reinterpretation of environmental health science data in an effort to delay regulatory reform or influence court cases and the general public.
- ☐ Have independent third party analyses.

Logistical Recommendations for NIEHS:

- ❑ Creation and long-term support of searchable data and sample repositories.
- ❑ Requirement for large collaborative projects to release data into a centralized web-based database.
- ❑ Develop specific guidelines based on the types of data sharing (voluntary and non-voluntary, etc.)

Computational Challenges:

- *“Analysis, not data creation, will be the fundamental hurdle preventing further advances in the field of Environmental Health”.*
- Massively parallel data analysis tools with data sharing networks and cloud (or grid) computing cyber-infrastructure will be emerging ideal systems to work towards.
- Data harmonization efforts need standardized measures of exposure from start of study. Consider where effort is best spent in analysis (pooling, meta-analysis, etc.).